

Análise de metodologias para identificação de cluster de acesso a documentação civil básica

Darlan Aragão Mesquita^{1,2}
Henrique Llacer Roig²

¹ Secretaria de Direitos Humanos da Presidência da República – SDH/PR
SCS Quadra 9, Lote C, Ed Parque Cidade Corporate, Torre A, 10º andar,
70308-200 – Brasília - DF, Brasil
darlan.mesquita@sedh.gov.br

² Universidade de Brasília –UnB/IG
Caixa Postal 04465 - 70.910-900 – Brasília - DF, Brasil
roig@unb.br

Abstract. The way indicators of access to basic human rights, guaranteed to children and adolescents by the federal constitution, may help the development of local government performance was studied by modeling a geodatabase with administrative record indicators produced by the Civil Registry Statistics of the IBGE and Live Birth Information System – SINASC of the Ministry of Health. The spatial autocorrelation between these indicators was analyzed using Moran's Autocorrelation Coefficient in 5 Brazilian States with varied shape and municipal structure in order to verify the possibility of standardization of methodology for conducting analysis. Different methods of conceptualization of spatial relationships were applied for each State. A high spatial autocorrelation for the indicators chosen was obtained, although this autocorrelation was apparently better explained by the contiguity of polygons than by distance. Clusters were formed independent of the Conceptualization of Spatial Relationships choose, but it was not possible to develop a methodology to define standards for definition of locals to combat child vulnerability.

Palavras-chave: spatial analysis, spatial statistics, socioeconomic indicators, GIS, Live birth information system – SINASC, Civil Registry Statistics, vulnerability child análise espacial, estatística espacial, indicadores socioeconômicos, geoprocessamento, SIG, SINASC, estatística de registro civil, vulnerabilidade infantil

1. Introdução:

Ao ser promulgada em 1988, a Constituição Federal trouxe ao Estado brasileiro obrigações que implicam uma série de serviços que devem ser prestados aos seus cidadãos, contudo para que um indivíduo possa ter acesso a esses benefícios é necessário que esse indivíduo seja reconhecido pelo Estado. No Brasil o registro civil de nascimento - RCN, a certidão de nascimento, é considerado o documento mais importante para acesso à cidadania, por ser o primeiro documento de reconhecimento de existência do cidadão e porque somente por meio dele é possível o reconhecimento civil para demais documentos. Apesar disso, nem todos os brasileiros têm acesso a esse direito. Existem no Brasil iniciativas para ampliar as possibilidades de acesso a esses serviços garantidos por direito pela constituição. Essas iniciativas visam garantir direitos econômicos, sociais, culturais e ambientais, contudo não poderão assegurar direitos civis e políticos. Portanto é essencial para o Estado brasileiro identificar possíveis agregações de locais onde existam pessoas sem registro para concentrar suas políticas de combate à exclusão social.

O registro civil de nascimento e vulnerabilidade infantil são amplamente difundidos em estudos sobre direitos humanos e cidadania. Apesar do conhecimento empírico que as diferenças regionais também se aplicam a essas questões, faltam estudos sobre a acessibilidade a direitos e definição de agregações geográficas de vulnerabilidade infantil. Para este estudo considera-se vulnerabilidade infantil a suscetibilidade de indivíduos de até 18 anos a privação de algum direito, seja por meio de trabalho, exploração sexual ou ausência de condições econômicas propícias.

Foram escolhidos cinco estados do “Compromisso mais Nordeste e Amazônia Legal pela Cidadania” para estudo de indicadores que possam embasar a vulnerabilidade infantil. Os estados foram escolhidos por apresentarem diversidade nos seguintes critérios: forma do estado; quantidade, densidade e variação de área de município; variação dos indicadores escolhidos. Assim, análises para identificação de autocorrelação espacial, por Moran I, e concentração de hot spot, por General G, foram realizadas para os municípios dos estados do Acre, Amazonas, Paraíba, Pernambuco e Piauí conforme Tabela 1.

Tabela 1. Comparação características dos estados utilizadas para escolha.

Estado	Forma do Estado	Características do Estado em relação aos Municípios		
		Quantidade	Densidade	Variação de tamanho/área
Acre	Pequeno alongado	Poucos	Baixa	Baixa
Amazonas	Grande quadrangular	Poucos	Baixa	Alta
Paraíba	Pequeno quadrangular	Muitos	Alta	Baixa
Pernambuco	Grande alongado	Muitos	Alta	Alta
Piauí	Grande alongado	Muitos	Alta	Baixa

1.1. Revisão da literatura

O acesso a bens e serviços públicos tem evoluído significativamente no Brasil. Como exemplo a redução significativa da quantidade de crianças que não são registradas no primeiro ano de nascida, o subregistro de nascimento. Estudos do IBGE mostram que no Brasil em 12 anos o subregistro passou de 30,2% para 8,2%. Esse aumento significativo de acesso ao registro ocasionado por um conjunto de ações governamentais torna cada vez mais difícil a identificação de possíveis locais onde esses serviços são mais necessários.

O registro de nascimento, realizado nos Cartórios, representa a oficialização da existência do indivíduo, de sua identificação e da sua relação com o Estado, condições fundamentais ao cidadão (Brasil, 2008). Por esse motivo desde a promulgação da constituição federal de 1988 tem-se criado ações para ampliação do acesso a certidão de nascimento.

A fim de mensurar o acesso o registro civil de nascimento, o IBGE calcula a taxa de subregistro, com base nas estimativas de nascimento por estado e a quantidade de registro. Estudos do IBGE sugerem a existência de uma dinâmica bastante diversa no território brasileiro, e sugere como principais fatores dessas diferenças as questões culturais, estas provavelmente influenciadas pelo acesso a serviços públicos, especialmente serviços de saúde e pela distância dos órgãos para registro (Brasil, 2008).

A análise em escala municipal da vulnerabilidade infantil é demanda recorrente das três esferas de governo, visto que as ações desses programas acontecem em escala local. Até 2008 as Estatísticas de Registro Civil forneciam indicadores nacionais e estaduais de subregistro, contudo em 2009, o IBGE lançou mapas do percentual de registros de nascimento realizados no prazo de até 90 dias por microrregião. Esse mapa avança ao desagregar informações estaduais, contudo ao agregar informações por microrregiões assume que o fenômeno de registro de nascimento está associado a essa divisão administrativa, o que pode não ser verdadeiro, podendo ser aprimoradas por meio de geoestatística ou análise espacial.

Mitchell (2005) destaca-se que não existe uma distância de análise correta. Recomenda que ao invés de pensar em distâncias deve-se pensar em um padrões para uma análise exploratória em múltiplas escalas e espaciais. Câmara (2005) e Mitchell (2005) sugerem a utilização de técnicas exploratórias de Moran e mapas de espalhamento de Moran para análises de indicadores sócias, visto que normalmente apresentam autocorrelação espacial global e local, dessa forma mostrariam agregações espaciais e permitiriam a indicação de áreas prioritárias para políticas públicas.

Métodos de análise global, como Moran e General G, foram adaptados para calcular a estatística de cada feição. Esses métodos foram desenvolvidos primeiramente para, além de identificar agrupamentos, observar como feições individuais contribuem para o padrão de valores identificados pelos métodos globais, bem como observar anomalias espaciais e outlier (Mitchell, 2005, p.165).

2. Metodologia de Trabalho

A malha municipal de 2005, escala de 1:500.000 em sistema geográfico, foi utilizada por ser a melhor escala disponível para todo Brasil, visto que a malha de 2007 está disponível apenas em 1:2.500.000. Os vetores foram obtidos no formato shapefile e armazenados em file geodatabase criada por ESRI® ArcCatalog 9.3. O arquivo original foi transformado para sistema projetado South America Albers Conic Equal Área, com ferramenta Project mantendo-se SAD69 como sistema geográfico. Polígonos pertencentes a um único município foram unidos para permitir realização de análises espaciais.

Tabelas de dados de infraestrutura e socioeconômicos: Os dados de infraestruturais e informações socioeconômicas foram armazenados e tabelas relacionadas em Microsoft Access 2007, formato de arquivo Microsoft Access 2002. As fontes das tabelas do banco foram registros administrativos disponibilizados pelo IBGE e Ministério da Saúde. Análises agregação de municípios com baixo percentual de Registro Civil de Nascimento - RCN no ano de nascimento da criança foi obtida por meio de consulta a tabela 2609 do Sistema IBGE de Recuperação Automática - SIDRA. Para análise do percentual de crianças que nasceram no período de 2004 a 2007 em domicílio, utilizou-se a ferramenta disponível pelo Datasus para acesso ao Sistema de Informações de Nascidos Vivos - Sinasc. Os nascimentos realizados nas categorias Hospitais e outros estabelecimentos de saúde foram agregados na categoria Estabelecimento de Saúde. Os nascimentos ocorridos nas categorias outros e ignorado foram excluídos da análise. Os estudos de RCN e proporção de nascimento em domicílio em relação a nascimentos em estabelecimentos de saúde foram feitos pela média dos três últimos anos. Para RCN, foi retirada a média por ano de atraso para realização do registro no período de 2005 a 2008, sendo utilizado para o cálculo o percentual da população registrada no primeiro ano em relação a registrada no mesmo período com idades de 1 a 6 anos, quando a criança chega a idade escolar.

Análises Espaciais: As ferramentas disponibilizadas no Arc toolbox para análise espacial de áreas foram utilizadas para medir distribuição geográfica, identificação de padrões e identificação de cluster.

Concepção de relação espacial: As análises de indicadores por estados foram realizadas segundo concepções dependentes de distância (inverso da distância, inverso do quadrado da distância inversa e distância fixa) e independentes (contigüidade de polígono, normalizado ou não normalizado). Para os modelos dependentes de distância foram testadas três distâncias euclidianas: a distância padrão estabelecida pelo software (quando o campo da distância limite é deixada em branco), menor distância para que todos os polígonos tenham ao menos um vizinho; cinquenta quilômetros, expectativa de deslocamento linear máximo para realizar registro ou parto; a distância de máxima aglomeração calculada pela ferramenta *Multi-distance Spatial Cluster Analysis (Ripleys K Function)*.

Mapamento de Clusters: Técnicas exploratórias de índice de Moran I foram utilizadas para definição de mapas de espalhamento construídos com Cluster and Outlier Analysis: Anselin Local Moran's I. Os modelos e distâncias escolhidas que apresentaram nível de significância de pelo menos 5% foram analisados, quando não houve pelo menos dois modelos que se adequasse a esse critério foram escolhidos modelos com nível de significância de 10%. Foram escolhidas para análises as distâncias com maior índice Moran I.

3. Resultados e Discussão

3.1. Análise estatística das áreas

Os nove estados da região Nordeste apresentam 1.792 municípios, três vezes mais municípios em uma área três vezes menor que a Amazônia Legal, representando uma densidade municipal por unidade de área nove vezes maior que a região Norte. Foi observada maior concentração de municípios com aproximadamente 5.000 km² (380) na Amazônia Legal, o Nordeste também apresenta grande concentração de municípios pequenos (866), contudo a primeira classe no Nordeste inclui municípios doze vezes menores, com até 400 km², que a Amazônia Legal.

Os estados escolhidos replicam a lógica de distribuição de estados global, tendo maior quantidade de municípios pequenos, contudo com diferenças bastante significativas quanto às classes. Essas diferenças de tamanho e forma dos municípios refletem a variação na distância entre centróides vizinho, no Nordeste as diferenças distâncias são, assim como as áreas, bem menores que o Norte. O estado da Paraíba, cuja forma é mais compacta que os demais da região, possui distribuição diferenciada dos demais, apesar de possuir concentração de municípios pequenos há maior distribuição entre as demais classes. O estado de Pernambuco, mais alongado, possui grande concentração municípios menores que 250 km², diferenciam-se dos demais por possuir em seu território o arquipélago de Fernando de Noronha, que, devido a distância, influência no cálculo da distância média do centróide, aumentando em 3,5 km a média do estado. A diferença de distâncias para inclusão dos polígonos vizinhos de 44,6km, excluindo Fernando de Noronha, para 498,2km, incluindo-o, refletirá nas análises automáticas que usarão o dado maior, mesmo que não represente a realidade do estado.

3.2. Mapeamento de Clusters:

Os resultados da análise espacial de agrupamento em múltiplas distâncias demonstram que os estados apresentaram curvas distintas entre si, contudo as curvas de indicadores para cada estado foram bastante semelhantes entre si. O agrupamento desses indicadores foi de aproximadamente em 190km para Amazonas, 127km para Piauí, 54km para Paraíba, 80km para registro civil de nascimento em Pernambuco e 135km para os demais indicadores de Pernambuco, segundo o modelo a distância ótima. As análises espaciais com limiares de distância não incluíram alguns municípios que não obtiveram nenhum vizinho. À distância de cinquenta quilômetros não foram contabilizados 55 no Amazonas, 16 municípios do Acre, 4 do Piauí e 1 de Pernambuco (Fernando de Noronha). Para melhor distância de cluster 4 municípios do Amazonas não tiveram vizinhos.

Distância de máximo agrupamento: O estudo de agrupamento em múltiplas distâncias apresentou pouca variação entre os indicadores estudados e a curva sem peso, indicando que os pesos entre as pontos talvez não sejam suficientemente distinto entre si ou que os pontos que apresentem a grande variações sejam poucos para distinguir a curva de peso padrão 1.

De forma geral todos os indicadores estaduais, exceção ao Acre, apresentaram aglomeração ao início nos primeiros quilometro diferenciado da reta esperada de aleatoriedade. O estado do Acre foi o único a não obter valores representativos de agrupamento para a análise de multidistâncias.

3.3. Nascimentos em domicílio

Na análise da proporção de nascidos em domicílio em relação a nascidos em estabelecimento de saúde, apenas Pernambuco obteve índice global com significância alta para todas as distâncias e todas as relações espaciais. A Paraíba não apresentou índices significância de 1%.

Todos os estados obtiveram relação de significância alta, pelo menos 5%, para o modelo de polígono contíguo. A distância de 50 km a distância apresentou a maior aleatoriedade.

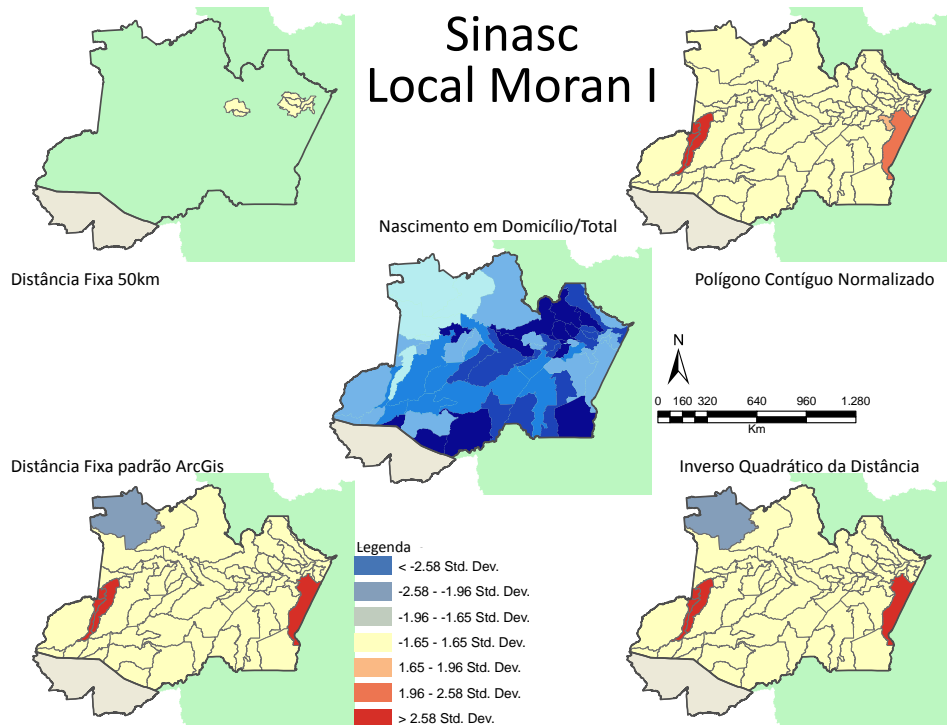


Figura 1. Análise de valores de escore Z do Índice de Moran I para percentual de nascimento em domicílio em diferentes concepções de relação espacial para Acre e Amazonas. Ao centro mapa de percentual de nascimento em domicílio com valores relativos utilizando pontos de quebra naturais para cada estado

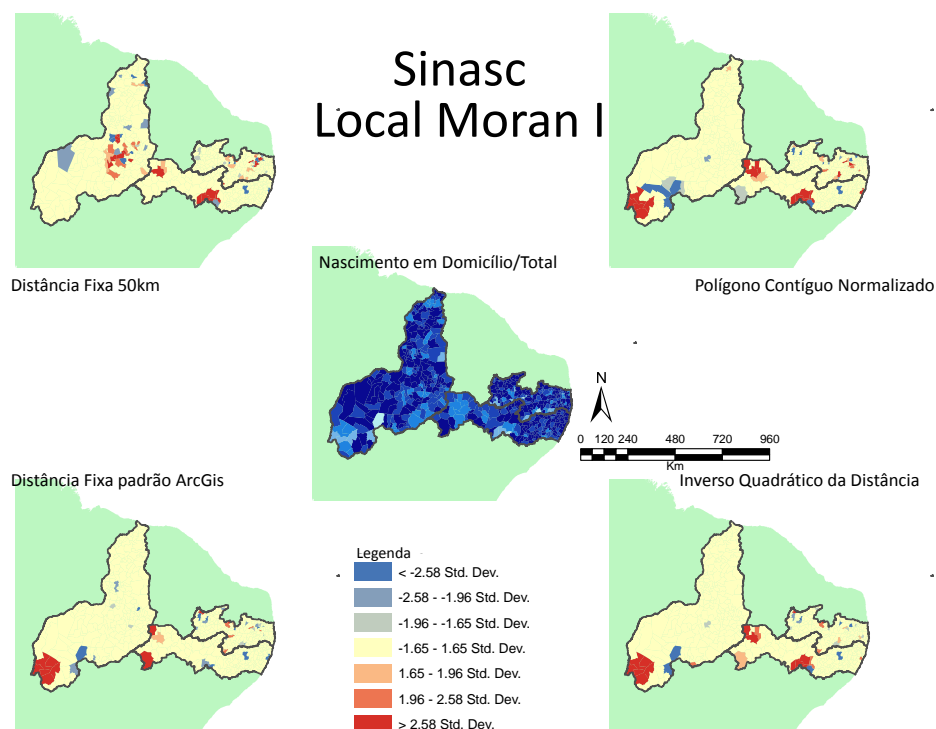


Figura 2. Análise de valores de escore Z do Índice de Moran I para percentual de nascimento em domicílio em diferentes concepções de relação espacial para Paraíba, Pernambuco e Piauí. Ao centro mapa de percentual de nascimento em domicílio com valores relativos utilizando pontos de quebra naturais para cada estado.

Todos os estados analisados, exceto o Acre, apresentaram os índices de Moran I maiores para modelos relacionado a vizinhança (polígono contíguo ou polígono contíguo normalizado) que os modelos que utilizam dependência de distância, sendo o índice para polígonos contínuos normalizados maior que para polígono contínuo. Modelos baseados em distância obtiveram agregação diferente por estado. Para os estados do Norte foram significativos principalmente modelos na qual todos os polígonos possuíam ao menos um vizinho, no caso do Acre apenas esses foi significativo. O Amazonas obteve índices maiores para esse tipo de escolha de distância, contudo também apresentou resultados significativos para melhor distância.

Os estados do Nordeste tiveram variações maiores. Todos os índices para o Piauí foram relativamente baixos e semelhantes, independente de modelo ou distância. A Paraíba apresentou índices significativos apenas para distâncias que incluíssem ao menos um vizinho. Pernambuco destaca-se por ter obtido índices maiores em distâncias limites de cinquenta quilômetros e os menores índices quanto a distância foi suficiente para incluir ao menos um município vizinho.

3.4. Registro Civil de Nascimento

O percentual de nascidos registrados no primeiro ano apresentou índices de significativa elevados para as análises. Das 52 análises, 44 modelagens obtiveram significância de 1%, 4 de 5%, 1 de 10% e 3 aleatórios (todos com distância de 50km para estados do Norte).

Semelhantemente aos índices Moran I obtidos para percentual de nascidos, os índices relacionado a vizinhança foram maiores que os modelos que utilizam dependência de distância, especialmente o de polígono contíguo normalizado, exceto pelo Piauí. Quanto a modelos dependentes de distância, em todos os casos, exceto Paraíba, por cinco centésimo, a relações espaciais de inverso quadrático da distância foram maiores que as demais.

Para o estado do Amazonas, o índice de Moran I foi de 3,6 a 6 vezes maior para a distância de cinquenta quilômetros, que apresentou significância de 5%, que os demais. Tendo por isso sido incluído nas análises locais.

3.5. Discussão: Nascimentos em domicílio, Registro Civil de Nascimento

Os indicadores analisados apesar de obter índices Moran I distintos apresentaram comportamentos semelhantes no que se refere a padrão de relação com distância e concepção de relacionamento espacial.

Conforme esperado, a pequena quantidade de município do estado do Acre e a não inclusão de 16 municípios aos cálculos tornou não significativa a análise na distância na qual todos os polígonos não possuísem ao menos um vizinho. Demonstrando procedência do aviso de que para análise são recomendadas ao menos 30 feições.

Ao deixar o software definir a distância limiar para Pernambuco sem considerar as interações reais, foi definida uma distância demasiadamente grande, a distância para qual todos os polígonos devem ter ao menos um vizinho. Apesar de garantir significância para as análises essa distância apresentou índices de Moran I abaixo dos demais.

Os mapas produzidos (Figuras 1 a 4) mostram coincidências em relação a área de agregação de agrupamentos e dispersão de nascimento. A análise dos diferentes tipos de concepções espaciais procurou identificar padrões de relacionamento espacial entre as variáveis estudadas. Para esses fenômenos sociais, obteve-se grande significância modelos que restringem a área de influência a municípios vizinhos.

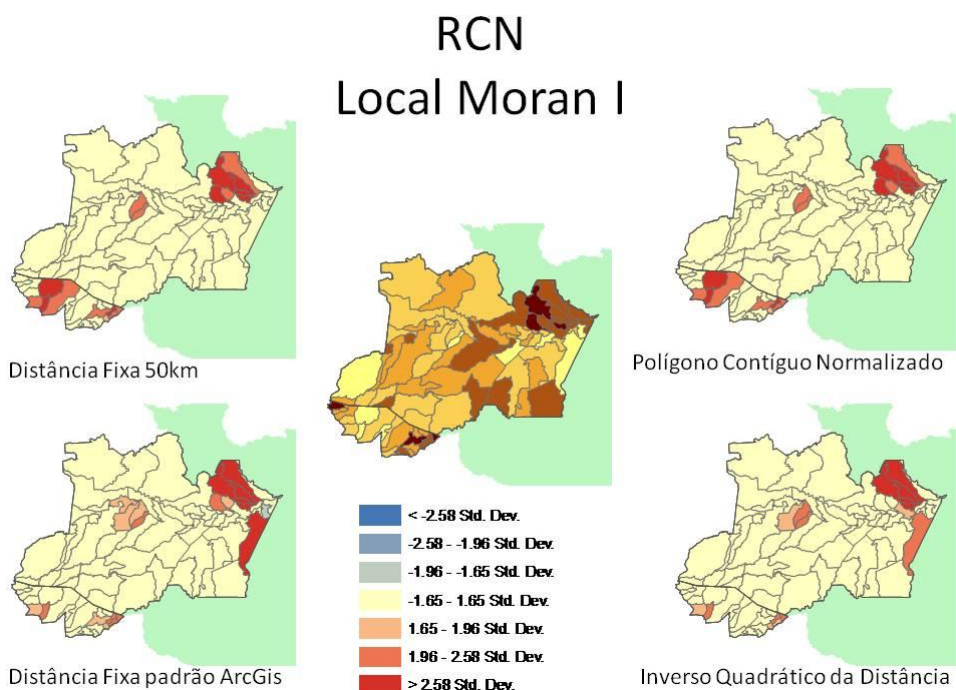


Figura 3. Análise de valores de escore Z do Índice de Moran I para Registro de Nascimento em diferentes concepções de relação espacial para Acre e Amazonas. Ao centro mapa de Registro de Nascimento em domicílio com valores relativos utilizando pontos de quebra naturais para cada estado.

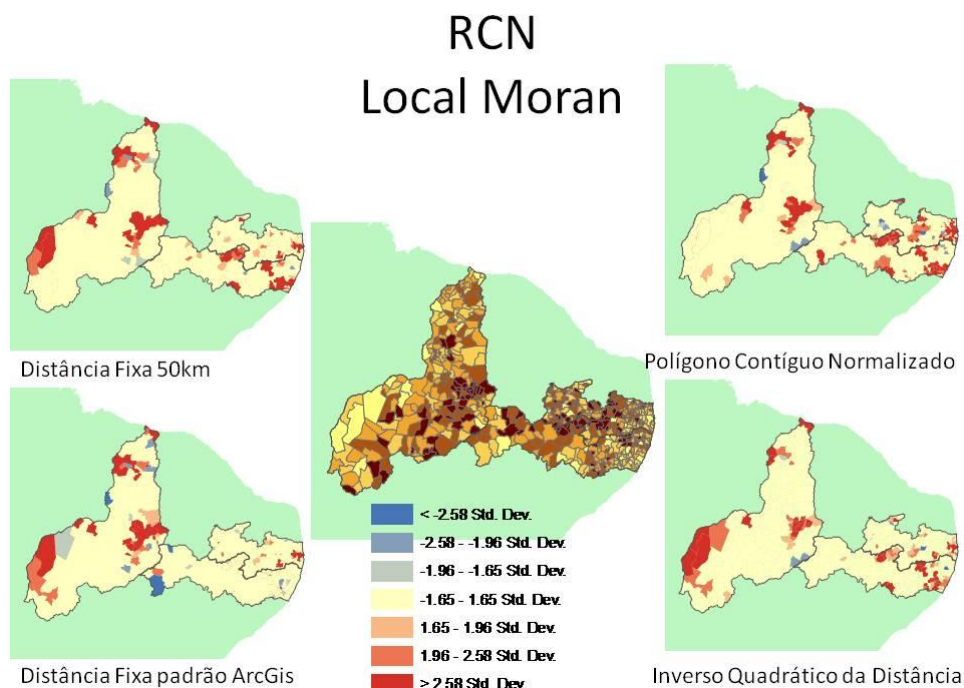


Figura 4. Análise de valores de escore Z do Índice de Moran I para Registro de Nascimento em diferentes concepções de relação espacial para Paraíba, Pernambuco e Piauí. Ao centro mapa de Registro de Nascimento em domicílio com valores relativos utilizando pontos de quebra naturais para cada estado.

Resta realizar análises mais detalhadas sobre a utilização de modelos que utilizam distância ou polígonos contíguos, visto que normalmente a literatura recomenda a utilização

de métodos de distâncias fixas para informações relacionadas a polígonos com grande variação de tamanho e de polígonos contíguos quando os polígonos apresentam tamanhos e distribuições similares. Nesse estudo os índices foram maiores quando a concepção de relação espacial considerou contiguidade. O que talvez indique a principal relação seja de proximidade de municípios.

Esse estudo corrobora que quando dados de distribuição não normal é aconselhável a utilização de distâncias que garantam a existência de ao menos um vizinho para cada polígono.

Destaca-se que apesar de existirem agrupamentos de locais com alto percentual de crianças que não são registradas antes de atingirem a idade escolar, a quantidade absoluta de pessoas sem registro pode ser maior em localidades com maior densidade populacional, apesar de que nesses locais os serviços públicos são mais acessíveis e constantemente demandados.

Considerando a semelhança no comportamento dos dados estudados, sugere-se para próximos estudos a realização de estudos de correlação entre os indicadores analisados, bem como a relação com variáveis demográficas.

A inclusão de quesito sobre a posse ou não de certidão de nascimento no Censo Demográfico 2010 permitirá verificar se as inferências desses modelos estudados sobre localização de subregistro de nascimento estão corretas.

4. Conclusões

Esse estudo atingiu parcialmente os seus objetivos visto que demonstrou existir autocorrelação espacial para os indicadores escolhidos. Não explorou a determinação exata de locais para atuação governamental, pois apesar de se verificar a formação de alguns padrões independentes do tipo de concepção espacial escolhida não foi possível desenvolver metodologia para definição padrão de locais para atuação local no combate a vulnerabilidade infantil, devido a diferença na organização e estrutura dos municípios.

Para análises futuras, a fim de verificar a adequação das concepções espaciais a cada realidade, serão realizadas simulações que considerem o efeito de borda, incluindo para isso região tampão. Como os fenômenos analisados aparentemente têm maior dependência da adjacência dos municípios que da distância, sugere-se que nos próximos estudos seja incluído modelo de concepção espacial de zona de indiferença.

Agradecimentos:

A Secretaria dos Direitos Humanos da Presidência da República pelo apoio ao projeto, à equipe do Instituto de Geociências da UnB e colegas da especialização pelas contribuições substantivas.

Referência bibliográfica

Brasil, Instituto Brasileiro de Geografia e Estatística - IBGE, **Estatísticas de Registro Civil**, Rio de Janeiro, v. 35 (2008), p.1- 170, 2009

Brasil, Instituto Brasileiro de Geografia e Estatística - IBGE, **Estudos e Pesquisas: Informação Demográfica e Socioeconômica. Indicadores Sociodemográficos e de Saúde no Brasil**, Rio de Janeiro, v 25 (2009)

Brasil, Ministério da Saúde. **Saúde Brasil 2007: uma análise da situação de saúde**. Brasília; Ministério da Saúde. 641 p. ilus, tab, graf; 2008

Câmara, G. Carvalho, M.S.; Cruz O. G., Correa, V. Análise de Dados de Área In:Druck, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.V.M. (eds) **Análise Espacial de Dados Geográficos**. Brasília, EMBRAPA, 2004

Mitchell, A. **The ESRI Guide to GIS Analysis: Volume 2: Spatial Measurements and Statistics** Redlands, ESRI Press, 2005.